School cost functions: a meta-regression analysis

By
Andrew D. Colegrave
and
Margaret J. Giles

Centre for Labour Market Research M409
The University of Western Australia
39 Myers Street
Crawley WA 6009
(email: mgiles@biz.uwa.edu.au)

Abstract:
The education cost literature includes econometric studies attempting to confirm or otherwise the need for class, school or school district consolidation. These studies date back to US state-specific analyses of school expenditure data undertaken by Riew (1966) in Wisconsin and Cohn (1968) in Iowa.

Many of these studies have attempted to determine economies of scale, or estimate an optimal school or district size. Not only do their results differ, but the studies use dissimilar data, techniques and models. To derive value from these studies requires that the estimates be made comparable. One method to do this is meta-regression analysis (MRA) which was pioneered by Jarrell and Stanley in the late 1980s as a result of similar frustrations with “omnipresent biases” (Stanley, 1998: 717).

In this paper, the basic technique of MRA is described and then applied to twenty-two estimates of school costs. Results suggest an optimal school size of around 1,543 students at the US secondary school level. Difficulties in interpreting the estimates are highlighted. In particular, the lack of consistency in the reporting across studies with similar aims and objectives is exposed. The paper concludes with cautions in the use of MRA and opportunities for further research in this area.

Paper to be presented at the 2005 Australian Conference of Economists, University of Melbourne, 26 – 28 September.
CONTENTS

Acknowledgements: ........................................................................................................ 3
I. Introduction ............................................................................................................. 4
II. Background ......................................................................................................... 5
III. Methodology ..................................................................................................... 6
   All relevant studies ........................................................................................... 6
   Parameter of interest ......................................................................................... 8
   Moderator (meta-independent) variables ..................................................... 9
   Estimation of the meta-regression model ..................................................... 10
   Specification tests .......................................................................................... 11
IV. MRA Results .................................................................................................... 12
V. Conclusions ...................................................................................................... 13
References ............................................................................................................. 15
Appendix: Meta-Independent Variables (Zk) ...................................................... 22

TABLES

Table 1: Optimal School Size and Meta-Independent Variables ................. 18
Table 2: Descriptive Statistics ........................................................................ 19
Table 3: Correlation Matrix ............................................................................. 20
Table 4: MRA Results: Dependent Variable is OSS .................................. 21
ACKNOWLEDGEMENTS:

The authors wish to thank Mr. Tom Stanley, Professor of Economics and Business, Hendrix College for his well received advice and encouraging comments, Professor Michael McAleer, The University of Western Australia, for his comments and support and econometrics advice, Mr. James Fogarty, The University of Western Australia, for advice on the conduct of an MRA and participants at the 2004/05 Australian Labour Market Research Workshop for helpful comments and anonymous referees. Naturally any errors and omissions are the authors.

This project has been funded by the Australian Research Council and the WA Department of Education and Training as a Linkage Project for 2003 – 2006.
I. INTRODUCTION

The education cost literature includes econometric studies attempting to confirm or otherwise the need for class, school or school district consolidation. These studies date back to specific state analyses of school expenditure data in the United States undertaken by Riew (1966) in Wisconsin and Cohn (1968) in Iowa. Both studies employed a standard ordinary least squares (OLS) regression technique with average total school cost per student as the dependent variable and school size in quadratic form as a predictor.

Since then, there have been other econometric studies using school level data and some studies using student or district data. Many of these studies use United States data at county or state level (for example, Bickel et al. (2000), Bowles and Bosworth (2002), Butler and Monk (1985), Callan and Santerre (1990), Chakraborty (2000), Daneshvary and Clauretie (2001), Deller and Rudnicki (1992), Duncombe, Miner and Ruggiero (1994; 1995), Duncombe, Ruggiero and Yinger (1996), Duncombe and Yinger (2003), Ferguson (1991), Gronberg et al. (circa 2004), Imazaki (2001), Imazaki and Reschovsky (2003), Monk et al. (1996), Osburn (1970), Reschovsky and Imazaki (2000), Riew (1986), and Stiefel et al. (1998)).

A number of econometric studies in Canada have also looked at school costs and the potential for economies of scale (for example, Dawson (1972), Kumar (1983) and Wales (1973)). Few studies have been undertaken in other jurisdictions (for example, Bee and Dolton (1985), Hough (1981) and Watt (1980) in the UK; Smet (2000) in Belgium and Jimenez (1986) in Bolivia and Paraguay). In particular, econometric studies of school costs using Australian data are rare (Hind, 1977).

Most of these studies have attempted to draw conclusions about economies of scale or estimate an optimal school or district size. Not only do all of these estimates differ, but the studies use dissimilar data, techniques and model specifications. For example, some studies have collected data on a school by school basis and others have observations at the district level. Some studies use OLS regression techniques and others use two-stage least squares to derive a total or average cost function(s). Furthermore, some studies have used log-linear models and others have used the more general functional form of the translog. It is reasonable to assume that some of the variation in parameter estimates (such as cost per student and optimal school or district size) is due to these and other differences in the conduct of the studies.

As Stanley and Jarrell (1989) question:

Have you ever wondered why there is so much variation among the reported empirical results of economic research? Why do researchers come to such different findings when they are purportedly investigating the same phenomenon? Does the reason lie in the idiosyncratic choices of statistical methods? Or, is it a result of the biases induced by model misspecifications? Perhaps it is the unique character of different data sets (pp. 161-2).

To derive value from these school cost studies requires that the estimates be made comparable. One method to do this is through meta-regression analysis
(MRA), which was pioneered by Jarrell and Stanley in the late 1980s as a result of similar frustrations with “omnipresent biases” (Stanley, 1998: 717).

In this paper, section II provides a background on MRA by introducing its predecessor meta-analysis. Section III details the five step methodology involved in conducting an MRA, highlighting the rigorous process involved in choosing the relevant studies, parameter of interest and meta-independent variables. Results of the application of MRA to twenty-two estimations of average school cost functions from ten studies are presented and discussed in section IV. The paper concludes with cautions in the use of MRA and suggestions for further research in this area.

II. BACKGROUND

Whilst Stanley and Jarrell (1989) recognize literature reviews undertaken by empirical economic researchers are instrumental in summarizing contending economic theories, they question the degree of subjectivity involved.

Traditionally, economists have not formally adopted any systematic or objective policy for dealing with the critical issues which surround literature surveys. As a result, reviews are rarely persuasive to those who do not already number among the converted (p. 162).

Thus they introduce, to an economic audience, MRA as a promising methodology of literature reviewing.

The background to the development of this methodology is from a popular technique called meta-analysis, used widely in the medical, psychology and education literature. Most of this literature is concerned with effect size, \( \omega = (\mu_e - \mu_c) / \sigma \), where \( \mu_e \) is the mean of one group (usually the experimental group), \( \mu_c \) is the mean of the control group, and \( \sigma \) is the standard deviation of the control group. Effect size, \( \omega \), is then used to compare parameter estimates from different studies. Its mathematical equivalence to Z-score transformation is not coincidental. In fact, the effect size is “a standard measure of empirical effect which can be assumed constant across the literature … this assumption … allows the meta-researcher to combine previous results and to investigate the process that generates these results” (Stanley and Jarrell, 1989: 163).

In this respect, meta-analysis is not terribly well suited to economics where the existence of control and experimental groups are rare. Thus meta-analysis has been reported, but rarely in the economics literature. The Journal of Economic Literature (which dates back to 1969 in electronic form and includes over 630,000 records) lists only 177 studies with either ‘meta analysis’ or ‘meta-analysis’ in its default fields. The most recent of these studies is a meta-analysis of fishing parameters (catch size, market values, etc) in Western Africa since 1960 (Alder and Sumailan, 2004). The earliest of the listed studies was called by the authors “a type of meta-analysis” and was a quantitative review of consumer behaviour research experiments published during the period 1970 to 1982 (Peterson, Albaum, and Beltramini, 1985).

One definition of meta analysis states that it is:
An overview in which quantitative methods are used to summarize the results of several studies on a single topic. A meta-analysis is used in an attempt to gain greater objectivity, generalizability and precision by including all the available high quality evidence from randomized controlled trials carried out on a specified topic. (Cambridge Institute for Research Education and Management (CiREM), n.d.).

In other words, meta-analysis takes a literature review to the next level by attempting to quantify effect size differences between studies. This is a valuable exercise but it still begs the question of what is the relative importance of the key factors within the study methodologies driving these differences. Thus enters MRA in the late 1980s, with Stanley and Jarrell's article in the Journal of Economic Surveys (1989).

The Journal of Economic Literature (JEL) lists only 13 studies with either ‘meta regression analysis’, ‘meta-regression analysis’, ‘meta regression’ or ‘meta-regression’ in its default fields. The earliest of these studies is the aforementioned paper by Stanley and Jarrell (1989), which provides a theoretical perspective on the usefulness of MRAs.

The advantages of MRA over meta-analysis are summarized succinctly by Stanley (1998):

Although it began as a method of combining and summarizing experimental findings, meta-regression analysis is designed explicitly to estimate and account for the omnipresent biases found in non experimental empirical economics … meta-regression analysis provides a means to estimate and thereby escape from these biases. While conventional narrative reviews may acknowledge the problems of empirical economic research, they are powerless to resolve them. (p. 717).

According to Stanley (2001), MRA requires five consecutive steps; discerning all relevant studies, choosing a summary statistic for the dependent variable, choosing moderator (meta-independent) variables related to study design and conduct, conducting the MRA, and conducting specification tests of the MRA. These steps will now be addressed.

III. METHODOLOGY

All relevant studies

The key features of studies included in this paper are that they examine economies of scale through analysis or estimation of school or school district cost functions using econometric techniques. The general methodology (summarised by Kumar, 1983) to examine economies of scale is to estimate average cost, $AC$, as a function of size, $S$, (such as enrolments), exogenous inputs, $X_1…X_n$, (such as institutional character, like school location and number of subjects offered), educational outcomes, $Q_1…Q_m$, (such as test scores or retention rates) and input prices, $p_1…p_k$, (such as teacher salaries). That is:

$$AC = f(S, X_1…X_n, Q_1…Q_m, p_1…p_k)$$  \hspace{1cm} (1)
for \( n \) exogenous inputs, \( m \) educational outcomes and \( k \) input prices.

In the presence of cross-sectional data, \( p_1\ldots p_k \) can be excluded as they can be assumed to be invariant across schools/districts within a single year. Thus equation (1) can be rewritten as:

\[
AC = f(S, X_1\ldots X_n, Q_1\ldots Q_m)
\]

The most common functional form for the estimation of equation (2) allows for the enrolment variable, \( S \), to be entered as a quadratic as follows:

\[
AC = \alpha_0 + \alpha_1 S + \alpha_2 S^2 + \sum_{i=1}^{n} \beta_i X_i + \sum_{j=1}^{m} \gamma_j Q_j + \varepsilon
\]

The error term, \( \varepsilon \), has the usual properties.

Average school costs (\( AC \)) are measured in a number of ways. Kumar (1983) uses total current expenditure per student. Other studies use variations of this, such as operating expenditures per student (Riew, 1966; 1986) or fees charged per student (Watt, 1980). Expenditure per student and cost per student will be used interchangeably in this paper (Riew, 1986, see footnote 6).

Initially, the school cost studies to be included in this MRA were identified from three sources. First, the review of the literature started with the seminal studies by Cohn (1968) in Iowa and Riew (1966) in Wisconsin. Both studies employed a standard OLS regression technique with the functional form of equation (3). Their results found the existence of significant economies of scale within the public high school systems of both Iowa and Wisconsin, using data from the early 1960’s. Using the Social Sciences Citation Index, all studies with these references were then examined for relevance. Further pertinent references cited in these studies were also scrutinized.

A second source of school cost studies was key international journals publishing in this field. These included all issues of the Journal of Education Finance, Education Economics and the Economics of Education Review. This process yielded a number of additional pertinent studies. Finally, an internet search (as at September 2004) using the Australian Google search engine (www.google.com.au) and search phrases such as ‘economies of scale’, ‘school cost’ and ‘optimum school size’ identified a number of reports and other documents.

As the first step in the MRA is to include all relevant studies, there are two issues here, one related to the word ‘all’ and the other concerning the word ‘relevant’. Although the literature search process was intended to be thorough, it is impossible to guarantee that ‘all’ studies have been discovered. For example, unpublished studies, studies that did not refer to Cohn (1968) and Riew (1966) and average school cost studies that did not publish in the Journal of Educational Finance, Education Economics or the Economics of Education Review will not have been included. A related issue raised by Stanley (2001) is that of publication bias, which eludes to the preference of journal editors publishing only those results that find statistically significant effects. "Among meta-analysts this is called the ‘file drawer’ problem, because researchers may be more likely to consign insignificant results to the ‘file drawer’ never to be published" (p. 146). Therefore, the specifications of the selected studies may well be prone to this publication bias.

In terms of ensuring the ‘relevance’ of the included studies, necessarily, the process requires that certain similarities exist between the reviewed studies. In this regard, the studies needed to show estimations of equation (2) or of the
A number of studies identified through the search process did not contain these estimations. These were Monk et al. (1996) providing only descriptive statistics, three studies providing only translog total cost functions (Butler and Monk (1985), Callan and Santerre (1990) and Jimenez (1986)), two studies providing only OLS total cost functions (Dawson (1972) and Hough (1981)), two studies using a hyperbolic $1/S$ term (Hind (1977) and Riew (1986)) and another study (Smet, 2000) using a rather different specification, incorporating fixed and marginal costs as the only regressors.

Most of the studies provided more than one set of results for the estimation of equation (3). Following Stanley and Jarrell (1998), multiple estimates from the same study were used as separate observations if they referred to different years. Estimates from dissimilar models using data from the same year were included as separate observations, as were estimates from similar models reported in separate publications by the same authors using the same data. Thus thirty studies, with fifty-one estimations were short-listed for the MRA.

However, not all of these estimations could be included as observations in the MRA. In particular, choosing the parameter of interest necessitated excluding just over half of these estimations.

**Parameter of interest**

The second step in the MRA process is to choose a parameter of interest. In studies of school costs, there are two parameters that might be usefully compared. These are optimal school size and minimum cost per student. Other parameters include the coefficient of the school size variable, which implies economies/diseconomies of scale, and goodness of fit measures.

Minimum cost per student had to be abandoned as a potential parameter of interest because it was only provided in two studies, with only eighteen of the thirty studies providing condescriptives. The coefficient of the school size variable was also discarded, because making this estimate comparable across studies was once again a problem due to the lack of means of variables; this meant logged models could not be converted to linear and vice versa. Furthermore, quadratic models, where an optimum school size can be calculated, are not comparable to that of linear models where there is no optimum. The use of a goodness of fit measure, such as R squared as the parameter of interest, has not received much support within MRA studies. Moreover, its usefulness as a measure of fundamental differences in school cost studies is dubious.

Only optimal school size (OSS) and optimal district size (ODS) appeared to be both interpretable as parameters of interest and be reported or able to be derived in many of the studies (forty of the fifty-one estimations in twenty-one of the thirty studies). As this literature review is a preamble to a study of school level costs in Western Australia, the fourteen studies with thirty-one estimates of OSS were further examined. Four of these were subsequently discarded for various reasons. Only ten studies (Bee and Dolton (1985), Cohn (1968), Daneshvary and Clauvertie (2001), Deller and Rudnicki (1992), Kumar (1983), Osburn (1970), Riew (1966; 1986), Wales (1973) and Watt (1980)), with twenty-two OSS estimates remained for inclusion in this MRA.

It is also pertinent to recognize that OSS, at least in theoretical expectations, is seen as a ‘size efficiency’ number that reflects how big a school
must be to exist, or how small it must get before it is closed. Thomas (1990) recognizes this approach as “helpful and worthwhile... in providing budgetary cost information of alternative patterns of school district organization” (p. 49). This is important because we are trying to determine what study characteristics, like models used and types of data, are causing the variation in what theory says should represent this ‘size efficiency’ number.

A related issue refers to whether the OSS calculated for each study is derived from a steep-sided or flat-bottomed average cost curve. A comparative and graphical analysis of each of the 22 average cost curves reveals that both types of curves occur. This suggests that for some studies an OSS is more likely to be defined as a range and for others a specific number. This is in line with existing studies of economies of scale. Three of these literature reviews are outlined below.

Fox (1981), the first comprehensive literature review of economies of size in education, draws together the key dimensions of some thirty studies. He finds an OSS for high schools in the range of 1400 to 1800 students. However, he makes an important caveat, stating that “the theoretical underpinnings of nearly all of the interpretable studies are deficient and some may suffer from data difficulties. As a result, though the direction of the results is clear, there are weaknesses in each study which raise doubt (sic) about the exact size of any economies” (p. 287). His conclusion thus supports a range rather than an exact value for OSS.

Two Australian studies refer to a target OSS. Lenahan (1983), in his review of economies of scale in schools and the nature of school costs, refers to the need to identify the right size for a ‘large’ school’ that on efficiency grounds becomes the basis for developing Target Recurrent Resource Standards for schools (p. 2). His findings consistently refer to larger schools being less costly to operate than smaller schools and that the gains from scale run out fairly rapidly in primary schools but not so for secondary schools. In terms of a ‘size efficiency’ number, he refers to the Australian primary school study by Hind (1977) in which economies of scale are exhausted at 100 students. He also refers to the OSS figures calculated for secondary schools by Cohn (1968) and Riew (1966) of around 1500 students. Importantly, Lenahan (1983) also refers to the research limitations of the reviewed studies and that generalizations may be quite inappropriate.

Finally, the second Australian study by McKenzie (1995), reviewed the existence of scale economies in Australian secondary schools and the international evidence for economies of scale in schools. He found in the UK that “for those schools where scale economies were identified, minimum costs occurred in the range from about 800 to 1200 students” (p. 118). He also refers to Riew’s (1986) OSS range of between 600 and 800 students. The findings for Australian government secondary schools were also around the 800 student mark. Analogous to the other literature reviews, there is reference to the heterogeneity of school and student types and different locations.

**Moderator (meta-independent) variables**

The third step in the MRA process is to choose the moderator, predictor, or meta-independent variables. These can be continuous variables, or binary variables reflecting the presence or absence of study characteristics. The binary
variables used in this study reflect the presence or absence of input quality variables (i.e. pupil teacher ratio, teacher quality and the percentage of students in special education), an output quality variable (i.e. pass rates or test scores), exogenous input variables (i.e. institutional characteristics), input price variables (i.e. instructional salaries), and variables depicting the type of data and study design (i.e. US study only, high school only, and the continuous variables, sample size and year of data).

Refer to the Appendix for more complete variable definitions and Table 1 for a list of the twenty-two OSS estimates together with their corresponding meta-independent variable study characteristics. Tables 2 and 3 contain the descriptive statistics and correlations respectively.

Table 2 shows a mean optimal school size of 1,558 students for the 22 estimations included and a mean sample size of 231 schools. About one third of these estimations included variables for instructional salaries and teacher quality and omitted variables related to institutional characteristics. One fifth of the estimations included a variable for the proportion of students in special education and about 60 percent excluded student outcome variables such as pass rates and test scores. Half of the estimations were derived from non US, primary/elementary school data and omitted a variable for pupil-teacher ratio.

Table 3 shows a number a high correlations, significant at the 0.01 level, like non US data with instructional salaries at around -67% suggesting that these variables may be substitutes. The correlation of omitted pupil teacher ratio and omitted institutional characteristics is around 68%. It appears that estimations that excluded a pupil-teacher ratio also excluded variables reflecting other institutional characteristics. Estimations that included the input quality variable teacher quality but omitted the output quality variable (pass rates or test scores) have a positive correlation of 43%, significant at the 0.05 level. It is possible that these variables are seen in these studies as quality variables (without specifying input or output) and hence may be treated as substitutes in the determination of average school costs.

**Estimation of the meta-regression model**

The MRA model suggested by Stanley and Jarrell (1989: 164) and estimated in the results section of this paper is:

\[ b_j = \beta + \sum_{k=1}^{K} \alpha_k Z_{jk} + \varepsilon_j \]

for \( j = 1, 2, \ldots, L \)

This equation can be interpreted as follows. The dependent variable \( b_j \) is the parameter of interest, which in this study is OSS, with the subscript \( j \) referring to the \( j^{th} \) study in \( L \) studies. In this study, \( L \) refers to the number of estimations (twenty-two) as some of the ten studies produced more than one result.

The constant term \( \beta \) represents the average OSS estimate predicted by the model when all of the dichotomous meta-independent and continuous variables are zero. Stanley and Jarrell (1998), further describe this as representing the study characteristics that may reasonably be regarded as “standard”, or consistent with the “best practice”. This choice, they say, “will always be a matter of professional judgment”. Espey and Thilmany (2000) analogously call it the 'base' model.
It is our judgment that researchers wanting to estimate an OSS should use independent variables derived from the functional form of equation (3). However, they should also include at least one input quality variable as well as an output quality variable. This choice and our type of data choices are also influenced by the descriptive statistics from Table 2. For our ‘standard’ model, when choosing between which three input quality variables and the two binary types of data variables to use, those with means of inclusion above 0.40, as shown in Table 2, were considered. For example, non US data, with a mean of 0.55, comprising study estimations from both the UK and Canada, also tells us that on average 0.45 (or 45%) of the estimations are from US studies. Consequently the US is used as our ‘standard’ country of data. Neither of the study estimations from the UK or Canada have means of inclusion above 0.40, therefore they were grouped together.

Our ‘standard’ model thus refers to high schools or secondary schools from the US and includes variables indicating pass rates or test scores, pupil teacher ratio and exogenous institutional characteristics. Binary predictor variables beginning with the word ‘omitted’, like omitted pass rates or test scores, are given a one (dummy) if a study excluded the variable or characteristic it refers to from their OSS estimation(s). Our ‘standard’ model, when this binary is set to zero, would therefore include the variable. The same explanation applies to binary predictors beginning with the word ‘non’. All other binary predictors, like % of students in special education, are given a one to indicate their presence in studies OSS estimation(s) relative to their exclusion in the ‘standard’ model.

The $Z_{jk}$ variables are the moderator variables that measure characteristics of the studies, such as the type of data used and variables omitted or included. The coefficients, $\alpha_k$, reflect the average bias in these study characteristics relative to our ‘standard’ model. The error terms, $\epsilon_j$, have the usual properties.

**Specification tests**

The final step in the MRA is to check that the assumptions underlying the estimation of the linear regression model in this paper are met. These assumptions and the tests of their validity are outlined in Gujarati (1995: 60-69).

The first two assumptions can be dealt with summarily. The model is linear in parameters and the $Z$ values are fixed in repeated sampling. Linearity of the model in MRA is assured as the coefficients of the binary predictors are interpreted as deviations from the ‘standard’ parameter value (the constant). The predictor values for the model presented in this paper are all binary.

Standard assumptions regarding the error term will be reviewed in the results section of this paper, such as tests for heteroscedasticity and multicollinearity, and a final assumption regarding model specification will also be reviewed.
IV. MRA RESULTS

Table 4 shows the results of the estimation of our MRA model from equation (4). This will now be discussed, followed by a summary of the specification tests. The model is estimated using OLS, has a sample size of twenty-two and the dependent variable is estimated optimal school size. This model is significant at the 1% level and has an adjusted R² of about 75%.

Specification tests showed that the Breusch-Pagan chi-square statistic for heteroscedasticity is insignificant, suggesting that the assumption of constant variance of the error terms holds. Multicollinearity tests showed that the zero-order correlation coefficient is less than 0.8 and that the ratio of the maximum and minimum eigenvalues (k = 126) excludes the presence of strong multicollinearity.

From this model, the constant, which represents our chosen ‘standard’ or ‘base’ model estimate of optimal school size, is about 1543 students. This is significant at the 1% level. It is also quite close to the mean OSS from Table 2 of around 1558 students.

The only output quality variable, omitted pass rates or test scores, significant at the 5% level, tells us that studies which excluded pass rates or test scores in their OSS estimation(s), on average negatively bias our ‘standard’ model estimate by some 504 students. Analogously, the input quality variable teacher quality, significant at the 5% level, tells us that studies which included teacher quality also reduce our base model estimation, this time at around 534 students. Remembering though, the difference here is that it refers to a study’s inclusion of this variable, rather than omitting it from their OSS estimation(s). The high positive correlation of around 43%, from Table 3, lends support to a degree of substitution between pass rates or test scores and teacher quality in studies OSS estimation(s).

The inclusion by a study of another input quality variable for the percentage of students in special education, although not significant, does have a high coefficient at around 385, indicating, on average, that these studies have a positive bias on our ‘standard’ model OSS estimation. Furthermore, our final input quality variable, omitted pupil teacher ratio, highly insignificant, tells us that studies which excluded pupil teacher ratio, on average, have a negligible impact on our ‘standard’ model.

The exogenous input variable, omitted institutional characteristics, with its positive and significant coefficient at the 10% level, tells us that studies which excluded institutional characteristics, on average, increase our ‘standard’ model estimation by some 506 students. From Table 3, the significant positive correlation between this and omitted pupil teacher ratio at around 68%, indicates a degree of uncertainty as to which of these omitted variables is causing the bias. Bee and Dolton (1985) highlight the potential similarities in these variables by including pupil teacher ratio as an exogenous input (institutional characteristic) variable and not as an input quality variable. However, because a number of other studies, for example; Kumar (1983) and Watt (1980) include pupil teacher ratio as an input quality variable and not as an exogenous input variable, we also did the same. The input price variable, instructional salaries, significant at the 1% level, has a large, average, positive bias on the ‘standard’ model estimation, of some 1,307 students.
Finally, both of the type of data variables, non US and non high school only are significant at the 5% and 1% levels respectively, with non US data having an average positive bias of some 658 students, and non high school only data having an average negative bias of some 786 students. Non US data’s impact may well be misleading because of its high negative correlations with both instructional salaries and omitted pupil teacher ratio at around -67% and -54% respectively. However, the coefficient of non high school only data does indeed make intuitive sense, because elementary or combined schools would be likely to cater for a smaller contingent of students. One would expect that high schools, with their larger curriculum and streaming of units would cater for a larger number of students. Wade (2004), further confirms this by referring to the ability of secondary school students to travel further, and as a result, secondary schools have a wider catchment area and are able to accommodate a greater number of students.

V. CONCLUSIONS

Meta-regression analysis (MRA) can be a useful tool in discerning differences between estimates of the same economic parameters. Its strength lies in its application of tried and tested regression techniques. Its weakness lies in its requirement that all relevant previous studies be included in the dataset. This requirement has proved to be a significant stumbling block for this current study.

Whilst it will never be known absolutely if all previous studies have been included, the short-list of thirty studies, with fifty-one estimations chosen as the baseline dataset, do have a similarity of purpose. They examine economies of scale through analysis or estimation of school or school district cost functions, using econometric techniques. However, relevance in MRA parlance is more than just a similarity of purpose, as with any econometric study, the dataset must contain enough observations with the same unit of measurement. Therefore, when choosing the dependent variable of interest, the process itself required a rigorous analysis of the thirty studies to determine which available parameter fitted the aforesaid criteria.

Furthermore, some studies reported their methodology and the limitations of their data thoroughly pursuant to interpreting their results; others were frustratingly vague in their reporting. Our intention here is not to criticize but rather to recognize that whilst scholarship generally applauds the uniqueness of studies (Florax et al. (2002) cited in footnote 4 in Garcia-Quevedo, 2004), it would have benefited the current study to have more uniform attention to detail across the previous studies. Subsequently, the sample used in our MRA was from ten studies providing twenty-two estimates of optimal school size (OSS), with some of these studies providing one or more estimates according to criteria defined in Stanley and Jarrell (1998).

Despite the various shortcomings outlined in the paper, the MRA results in the current study have highlighted a number of impacts of estimates of optimal school size that need to be considered in studies of school costs. In our model, the ‘standard’ optimal school size is about 1543 students. Significant negative biases from this ‘standard’ were the omission of a variable denoting pass rates or test scores in previous studies, the inclusion of a variable denoting teacher quality and the inclusion of a variable denoting non high school only data. Significant positive biases were the omission of a variable depicting institutional
character, the inclusion of an instructional salaries variable, as well as the variable illustrating those studies that used data from outside the US. Correlations between some of these variables were quite high, so a degree of caution would have to be taken with reference to these results.

This preferred model explains just over seventy-five percent of the variation in optimal school size estimates. This is comparable to other MRAs. For example, in the thirteen meta-regressions published in the Journal of Economic Literature between 1989 and 2004, the $R^2$ or adjusted $R^2$ measure ranged from 0.027 (Model 2 in Garcia-Quevedo, 2004) to 0.84 (Doucouliagos and Laroche, 2003). The median value $R^2$ from these thirteen studies is 0.53.

Undertaking an MRA of studies of school cost functions was not straightforward and the interpretation of the results is far from clear. However, there are two implications of this process. First, it appears unequivocal that estimated school cost functions are peculiar to the jurisdictions from which the data are collected and the school system within which the schools operate. This was highlighted by the literature review studies of school cost functions, which demonstrated that wide variations exist for some theoretically expected ‘size efficiency’ number or range for an OSS. Lenahan (1983) aptly sums up here, by stating that “research into cost functions labours under many of the conceptual and measurement difficulties that bedevil enquiry into the social sciences generally, and the field of education in particular” (p. 13). Legislators, school boards and education ministries should therefore look at the cost structure and costs within their own systems and schools to optimize school or district size. Extrapolating the results from other systems and schools is unlikely to bring about the best (minimum cost) result.

Second, future econometric studies of school cost functions should accommodate the possibility of diminishing returns by including enrolments in quadratic form. They should also continue to include our ‘standard’ model variables, but perhaps with the inclusion of an additional input quality variable to indicate teacher quality, as well as consideration given to the inclusion of the input price variable instructional salaries. Importantly, reporting of the results of such studies should be mindful of minimum reporting requirements highlighted in this paper.

As a final remark, conducting future MRA’s in non-experimental empirical economics, although not without its challenges, certainly promotes a thorough investigation of the available literature and we must not lose sight of this potential benefit as a supplement to traditional methods of literature reviewing.
REFERENCES


## Table 1: Optimal School Size and Meta-Independent Variables

<table>
<thead>
<tr>
<th>Optimal school size</th>
<th>Omitted pass rates or test scores</th>
<th>% of students in special education</th>
<th>Non US data</th>
<th>Non high school only data</th>
<th>Instructional salaries</th>
<th>Omitted pupil-teacher ratio</th>
<th>Teacher quality</th>
<th>Omitted institutional characteristics</th>
<th>Sample size</th>
<th>Year of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1865</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>309</td>
<td>1980.5</td>
</tr>
<tr>
<td>2325</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>309</td>
<td>1980.5</td>
</tr>
<tr>
<td>2155</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>309</td>
<td>1980.5</td>
</tr>
<tr>
<td>2440</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>309</td>
<td>1980.5</td>
</tr>
<tr>
<td>2210</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>377</td>
<td>1962.5</td>
</tr>
<tr>
<td>1653</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>377</td>
<td>1962.5</td>
</tr>
<tr>
<td>1483</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>377</td>
<td>1962.5</td>
</tr>
<tr>
<td>774</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>115</td>
<td>1994.5</td>
</tr>
<tr>
<td>1544</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>35</td>
<td>1977.0</td>
</tr>
<tr>
<td>935</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>38</td>
<td>1976.0</td>
</tr>
<tr>
<td>1686</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>37</td>
<td>1975.0</td>
</tr>
<tr>
<td>2244</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>433</td>
<td>1966.0</td>
</tr>
<tr>
<td>1675</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>109</td>
<td>1960.5</td>
</tr>
<tr>
<td>1025</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>29</td>
<td>1978.5</td>
</tr>
<tr>
<td>1055</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>29</td>
<td>1978.5</td>
</tr>
<tr>
<td>1539</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>242</td>
<td>1969.0</td>
</tr>
<tr>
<td>514</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1211</td>
<td>1969.0</td>
</tr>
<tr>
<td>886</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>41</td>
<td>1968.0</td>
</tr>
<tr>
<td>755</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>1968.0</td>
</tr>
<tr>
<td>524</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>74</td>
<td>1968.0</td>
</tr>
<tr>
<td>2075</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>147</td>
<td>1987.0</td>
</tr>
<tr>
<td>2917</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>147</td>
<td>1987.0</td>
</tr>
</tbody>
</table>

Notes: The data contain 22 observations. (a) Year of data ending in .5 indicates the midpoint for data obtained between two years (i.e. 1980/81 = 1980.5). (b) Indicates the midpoint for the year of data range between 1985 and 1989.
Table 2: Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal School Size</td>
<td>1558.14</td>
<td>678.492</td>
</tr>
<tr>
<td>Omitted pass rates or test scores</td>
<td>.59</td>
<td>.503</td>
</tr>
<tr>
<td>% of students in special education</td>
<td>.18</td>
<td>.395</td>
</tr>
<tr>
<td>Non US data</td>
<td>.55</td>
<td>.510</td>
</tr>
<tr>
<td>Non high school only data</td>
<td>.50</td>
<td>.512</td>
</tr>
<tr>
<td>Instructional salaries</td>
<td>.27</td>
<td>.456</td>
</tr>
<tr>
<td>Omitted pupil teacher ratio</td>
<td>.50</td>
<td>.512</td>
</tr>
<tr>
<td>Teacher quality</td>
<td>.36</td>
<td>.492</td>
</tr>
<tr>
<td>Omitted institutional characteristics</td>
<td>.32</td>
<td>.477</td>
</tr>
<tr>
<td>Sample size</td>
<td>230.82</td>
<td>261.002</td>
</tr>
<tr>
<td>Year of data</td>
<td>1974.16</td>
<td>9.212</td>
</tr>
<tr>
<td>No. of observations</td>
<td></td>
<td>22</td>
</tr>
</tbody>
</table>
### Table 3: Correlation Matrix

<table>
<thead>
<tr>
<th>Variable</th>
<th>Optimal school size</th>
<th>Omitted pass rates or test scores</th>
<th>% of students in special education</th>
<th>Non US data</th>
<th>Non high school only data</th>
<th>Instructional salaries</th>
<th>Omitted pupil teacher ratio</th>
<th>Teacher quality</th>
<th>Omitted institutional characteristics</th>
<th>Sample size</th>
<th>Year of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal school size</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Omitted pass rates or test scores</td>
<td>-.695*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of students in special education</td>
<td>-.23</td>
<td>.153</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non US data</td>
<td>-.211</td>
<td>.169</td>
<td>.194</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non high school only data</td>
<td>-.313</td>
<td>.092</td>
<td>.471**</td>
<td>.183</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructional salaries</td>
<td>.527*</td>
<td>-.321</td>
<td>-.289</td>
<td>-.671*</td>
<td>.000</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Omitted pupil teacher ratio</td>
<td>.043</td>
<td>-.092</td>
<td>-.236</td>
<td>-.548*</td>
<td>-.091</td>
<td>.204</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher quality</td>
<td>-.291</td>
<td>.437**</td>
<td>-.356</td>
<td>-.259</td>
<td>-.378</td>
<td>.174</td>
<td>.189</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Omitted institutional characteristics</td>
<td>-.044</td>
<td>.171</td>
<td>-.322</td>
<td>-.356</td>
<td>-.098</td>
<td>.020</td>
<td>.683*</td>
<td>.295</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>.026</td>
<td>-.137</td>
<td>-.323</td>
<td>.060</td>
<td>-.084</td>
<td>.082</td>
<td>.217</td>
<td>.204</td>
<td>.217</td>
<td>.217</td>
<td>1.000</td>
</tr>
<tr>
<td>Year of data</td>
<td>.188</td>
<td>-.530**</td>
<td>.339</td>
<td>.021</td>
<td>.199</td>
<td>-.221</td>
<td>.174</td>
<td>-.470</td>
<td>.134</td>
<td>-.241</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Notes: *, and ** denote significance at the 0.01 and 0.05 levels.
### Table 4: MRA Results: Dependent Variable is OSS

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>t-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>1543.510</td>
<td>5.215*</td>
</tr>
<tr>
<td>Omitted pass rates or test scores</td>
<td>-504.077</td>
<td>-2.448**</td>
</tr>
<tr>
<td>% of students in special education</td>
<td>385.322</td>
<td>1.496</td>
</tr>
<tr>
<td>Non US data</td>
<td>658.137</td>
<td>2.527**</td>
</tr>
<tr>
<td>Non high school only data</td>
<td>-786.304</td>
<td>-4.078*</td>
</tr>
<tr>
<td>Instructional salaries</td>
<td>1307.381</td>
<td>4.687*</td>
</tr>
<tr>
<td>Omitted pupil teacher ratio</td>
<td>-93.293</td>
<td>-.408</td>
</tr>
<tr>
<td>Teacher quality</td>
<td>-534.403</td>
<td>-2.463**</td>
</tr>
<tr>
<td>Omitted institutional characteristics</td>
<td>506.051</td>
<td>2.099***</td>
</tr>
<tr>
<td>Adjusted R(^2)</td>
<td>0.753</td>
<td></td>
</tr>
<tr>
<td>F-test</td>
<td>9.009*</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>

Note: *, **, *** denote significance at the 0.01, 0.05 and 0.10 levels, respectively.
(1) OUTPUT QUALITY:
Omitted pass rates or test scores = 1 if a study omitted a variable(s) to indicate pass rates or test scores.

Bee and Dolton (1985) use average number of passes per pupil in General Certificate of Education (GCE) advanced level, which captures the general extent of educational qualifications achieved by the average sixth-former. Cohn (1968) uses the gain in academic achievements by high school pupils, calculated as the difference between the average composite score on the ITED (Iowa Tests of Educational Development) for the twelfth grade and the tenth grade. Daneshvary and Clauretie (2001) uses the percentage of fourth grade students scoring above the national average in (combined) math, language and reading. Deller and Rudnicki (1992) use the cumulative three year (1986-1989) Maine achievement test scores. Osburn (1970) uses the Ohio psychological test score.

(2) INPUT QUALITY:
Teacher quality = 1 if a study included a variable(s) to indicate teacher quality.

Cohn (1968) uses the average number of different subject matter assignments per high school teacher. Riew (1966) uses the average number of courses taught per teacher. Riew (1986) uses two measures; the first represents training, measured by teachers with a masters degree, as a percentage of total classroom teachers. The second represents teacher experience, measured by the number of teachers with seven years or more of classroom experience as a percentage of all classroom teachers. Wales (1973) uses two measures, being the average number of years teaching experience and the average number of years of education of the teachers. Watt (1980) uses the proportion of teachers with an MA degree.

Omitted pupil teacher ratio = 1 if a study omitted a pupil teacher ratio.

Percentage of students in special education = 1 if a study included a variable to indicate the % of students in special education.

(3) INPUT PRICES:
Instructional salaries = 1 if a study includes a variable for instructional salaries.

(4) EXOGENOUS INPUTS:
Omitted institutional characteristics = 1 if a study omitted a variable(s) to indicate institutional characteristics.

Bee and Dolton (1985) use a number of measures. First of all, they use sets of dummies to indicate the sex composition of the pupils; whether the school is a boarding or day school; and also for the schools geographical location. Then secondly, they use more specific measures, like the range of sub-advanced level courses provided by the school, indicated by the number of GCE Ordinary level languages and the number of Certificate of Secondary Education (CSE) subjects offered; and the links with the state sector, measured as the percentage of pupils originating from the state primary schools. Cohn (1968) uses the number of credit units offered for a full school year. Daneshvary and Clauretie (2001) include a dummy variable equal to one for schools with a year-round schedule and zero for schools with traditional schedules. Kumar (1983) includes two dummies to indicate the area the school is located in. Osburn (1970) uses two measures. The first, school classification, uses a set of three dummy variables indicating a AAA, AA or unclassified school, with schools classified to A specified as zero. These variables reflect the breadth of curriculum or number of credit units offered. The second, geographical location, also a dummy variable, is set to one if the school district is located in the Ozark region, to indicate differences with respect to expenditures and quality.
Riew (1966) uses two measures. The first represents the breadth of the curriculum, by using the number of credit units offered (a two-semester course meeting five times a week is counted as one unit) and the second, the percentage of classrooms built after 1950 to reflect variations of maintenance and operations costs by the ages of the schools properties. Watt (1980) includes two measures. The first represents the age of the school shown by the date of foundation and the second, the proportion of students financed by the local authority, which is likely to indicate the degree of control a local authority exerts over a school.

(5) TYPE OF DATA AND STUDY DESIGN:
Non high school only data = 1 if a study estimate does not use high school only data (which includes secondary and middle schools).
This refers to elementary school only estimates, as in Daneshvary and Clauverture (2001); Deller and Rudnicki (1992) and Wales (1973); and combined high school and elementary school estimates, as in Kumar (1983); Osburn (1970); and Watt (1980).
Non US data = 1 if a study estimate did not use data from the United States (US).
Bee and Dolton (1985) and Watt (1980) are from the UK; Kumar (1983) and Wales (1973) are from Canada.
Sample size = The sample size for each study estimate (continuous variable).
Year of data = The year of data for each study estimate (continuous variable).